A Simple Sample Size Formula for Estimating Means of Poisson Random Variables *

Xinjia Chen

Submitted in April, 2008

Abstract

In this paper, we derive an explicit sample size formula based a mixed criterion of absolute and relative errors for estimating means of Poisson random variables.

1 Sample Size Formula

It is a frequent problem to estimate the mean value of a Poisson random variable based on sampling. Specifically, let X be a Poisson random variable with mean $\mathbb{E}[X] = \lambda > 0$, one wishes to estimate λ as

$$\widehat{\boldsymbol{\lambda}} = \frac{\sum_{i=1}^{n} X_i}{n}$$

where $X_1, \cdots, X_n$ are i.i.d. random samples of X. Since $\widehat{\boldsymbol{\lambda}}$ is of random nature, it is important to control the statistical error of the estimate. For this purpose, we have

Theorem 1 *Let $\varepsilon_a > 0,\ \varepsilon_r \in (0,1)$ and $\delta \in (0,1)$. Then*

$$\Pr \left\{ \left| \widehat{\boldsymbol{\lambda}} - \lambda \right| < \varepsilon_a \text{ or } \left| \widehat{\boldsymbol{\lambda}} - \lambda \right| < \varepsilon_r \lambda \right\} > 1 - \delta$$

provided that

$$n > \frac{\varepsilon_r}{\varepsilon_a} \times \frac{\ln \frac{2}{\delta}}{(1+\varepsilon_r)\ln(1+\varepsilon_r) - \varepsilon_r}. \tag{1}$$

It should be noted that conventional methods for determining sample sizes are based on normal approximation, see [3] and the references therein. In contrast, Theorem 1 offers a rigorous method for determining sample sizes. To reduce conservatism, a numerical approach has been developed by Chen [1] which permits exact computation of the minimum sample size.

*The author is currently with Department of Electrical Engineering, Louisiana State University at Baton Rouge, LA 70803, USA, and Department of Electrical Engineering, Southern University and A&M College, Baton Rouge, LA 70813, USA; Email: chenxinjia@gmail.com

2 Proof of Theorem 1

We need some preliminary results.

Lemma 1 *Let K be a Poisson random variable with mean $\theta > 0$. Then,* $\Pr\{K \geq r\} \leq e^{-\theta} \left(\frac{\theta e}{r}\right)^r$ *for any real number $r > \theta$ and* $\Pr\{K \leq r\} \leq e^{-\theta} \left(\frac{\theta e}{r}\right)^r$ *for any positive real number $r < \theta$.*

Proof. For any real number $r > \theta$, using the Chernoff bound [2], we have

$$\begin{aligned}\Pr\{K \geq r\} &\leq \inf_{t>0} \mathbb{E}\left[e^{t(K-r)}\right] = \inf_{t>0} \sum_{i=0}^{\infty} e^{t(i-r)} \frac{\theta^i}{i!} e^{-\theta} \\ &= \inf_{t>0} e^{\theta e^t} e^{-\theta} e^{-r\,t} \sum_{i=0}^{\infty} \frac{(\theta e^t)^i}{i!} e^{-\theta e^t} = \inf_{t>0} e^{-\theta} e^{\theta e^t - r\,t},\end{aligned}$$

where the infimum is achieved at $t = \ln\left(\frac{r}{\theta}\right) > 0$. For this value of t, we have $e^{-\theta} e^{\theta e^t - tr} = e^{-\theta}\left(\frac{\theta e}{r}\right)^r$. It follows that $\Pr\{K \geq r\} \leq e^{-\theta}\left(\frac{\theta e}{r}\right)^r$ for any real number $r > \theta$.

Similarly, for any real number $r < \theta$, we have $\Pr\{K \leq r\} \leq e^{-\theta}\left(\frac{\theta e}{r}\right)^r$.

□

In the sequel, we shall introduce the following function

$$g(\varepsilon, \lambda) = \varepsilon + (\lambda + \varepsilon) \ln \frac{\lambda}{\lambda + \varepsilon}.$$

Lemma 2 *Let $\lambda > \varepsilon > 0$. Then,* $\Pr\left\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon\right\} \leq \exp\left(n\, g(-\varepsilon, \lambda)\right)$ *and $g(-\varepsilon, \lambda)$ is monotonically increasing with respect to $\lambda \in (\varepsilon, \infty)$.*

Proof. Letting $K = \sum_{i=1}^n X_i$, $\theta = n\lambda$ and $r = n(\lambda - \varepsilon)$ and applying Lemma 1, for $\lambda > \varepsilon > 0$, we have

$$\Pr\left\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon\right\} = \Pr\{K \leq r\} \leq e^{-\theta}\left(\frac{\theta e}{r}\right)^r = \exp\left(n\, g(-\varepsilon, \lambda)\right),$$

where $g(-\varepsilon, \lambda)$ is monotonically increasing with respect to $\lambda \in (\varepsilon, \infty)$ because

$$\frac{\partial g(-\varepsilon, \lambda)}{\partial \lambda} = -\ln\left(1 - \frac{\varepsilon}{\lambda}\right) - \frac{\varepsilon}{\lambda} > 0$$

for $\lambda > \varepsilon > 0$.

□

Lemma 3 *Let $\varepsilon > 0$. Then,* $\Pr\left\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon\right\} \leq \exp\left(n\, g(\varepsilon, \lambda)\right)$ *and $g(\varepsilon, \lambda)$ is monotonically increasing with respect to $\lambda \in (0, \infty)$.*

Proof. Letting $K = \sum_{i=1}^n X_i$, $\theta = n\lambda$ and $r = n(\lambda + \varepsilon)$ and applying Lemma 1, for $\lambda > 0$, we have

$$\Pr\left\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon\right\} = \Pr\{K \geq r\} \leq e^{-\theta}\left(\frac{\theta e}{r}\right)^r \leq \exp\left(n\, g(\varepsilon, \lambda)\right),$$

where $g(\varepsilon, \lambda)$ is monotonically increasing with respect to $\lambda \in (0, \infty)$ because

$$\frac{\partial g(\varepsilon, \lambda)}{\partial \lambda} = -\ln\left(1 + \frac{\varepsilon}{\lambda}\right) + \frac{\varepsilon}{\lambda} > 0.$$

□

Lemma 4 *$g(\varepsilon, \lambda) > g(-\varepsilon, \lambda)$ for $\lambda > \varepsilon > 0$.*

Proof. Since $g(\varepsilon, \lambda) - g(-\varepsilon, \lambda) = 0$ for $\varepsilon = 0$ and

$$\frac{\partial\left[g(\varepsilon, \lambda) - g(-\varepsilon, \lambda)\right]}{\partial \varepsilon} = \ln\frac{\lambda^2}{\lambda^2 - \varepsilon^2} > 0$$

for $\lambda > \varepsilon > 0$, we have

$$g(\varepsilon, \lambda) - g(-\varepsilon, \lambda) > 0$$

for any $\varepsilon \in (0, \lambda)$. Since such arguments hold for arbitrary $\lambda > 0$, we can conclude that

$$g(\varepsilon, \lambda) > g(-\varepsilon, \lambda)$$

for $\lambda > \varepsilon > 0$. □

Lemma 5 *Let $0 < \varepsilon < 1$. Then, $\Pr\left\{\widehat{\boldsymbol{\lambda}} \leq \lambda(1 - \varepsilon)\right\} \leq \exp\left(n\, g(-\varepsilon\lambda, \lambda)\right)$ and $g(-\varepsilon\lambda, \lambda)$ is monotonically decreasing with respect to $\lambda > 0$.*

Proof. Letting $K = \sum_{i=1}^n X_i$, $\theta = n\lambda$ and $r = n\lambda(1 - \varepsilon)$ and making use of Lemma 1, for $0 < \varepsilon < 1$, we have

$$\Pr\left\{\widehat{\boldsymbol{\lambda}} \leq \lambda(1 - \varepsilon)\right\} = \Pr\{K \leq r\} \leq e^{-\theta}\left(\frac{\theta e}{r}\right)^r \leq \exp\left(n\, g(-\varepsilon\lambda, \lambda)\right),$$

where

$$g(-\varepsilon\lambda, \lambda) = \left[-\varepsilon - (1 - \varepsilon)\ln(1 - \varepsilon)\right]\lambda,$$

which is monotonically decreasing with respect to $\lambda > 0$, since $-\varepsilon - (1 - \varepsilon)\ln(1 - \varepsilon) < 0$ for $0 < \varepsilon < 1$.

□

Lemma 6 *Let $\varepsilon > 0$. Then, $\Pr\left\{\widehat{\boldsymbol{\lambda}} \geq \lambda(1 + \varepsilon)\right\} \leq \exp\left(n\, g(\varepsilon\lambda, \lambda)\right)$ and $g(\varepsilon\lambda, \lambda)$ is monotonically decreasing with respect to $\lambda > 0$.*

Proof. Letting $K = \sum_{i=1}^n X_i,\ \theta = n\lambda$ and $r = n\lambda(1+\varepsilon)$ and making use of Lemma 1, for $\varepsilon > 0$, we have

$$\Pr\left\{\widehat{\boldsymbol{\lambda}} \geq \lambda(1+\varepsilon)\right\} \quad \leq \quad \exp\left(n\ g(\varepsilon\lambda, \lambda)\right)$$

where

$$g(\varepsilon\lambda, \lambda) = [\varepsilon - (1+\varepsilon)\ln(1+\varepsilon)]\,\lambda,$$

which is monotonically decreasing with respect to $\lambda > 0$, since $\varepsilon - (1+\varepsilon)\ln(1+\varepsilon) < 0$ for $\varepsilon > 0$.

□

We are now in a position to prove the theorem. It suffices to show

$$\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r\lambda\right\} < \delta$$

for n satisfying (1). It can shown that (1) is equivalent to

$$\exp(n\ g(\varepsilon_a, \varepsilon_a)) < \frac{\delta}{2}. \tag{2}$$

We shall consider four cases as follows.

Case (i): $0 < \lambda < \varepsilon_a$;

Case (ii): $\lambda = \varepsilon_a$;

Case (iii): $\varepsilon_a < \lambda \leq \frac{\varepsilon_a}{\varepsilon_r}$;

Case (iv): $\lambda > \frac{\varepsilon_a}{\varepsilon_r}$.

In Case (i), we have $\Pr\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon_a\} = 0$ and

$$\begin{aligned}\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r\lambda\right\} &= \Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a\right\} \\ &= \Pr\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon_a\} + \Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\} \\ &= \Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\}.\end{aligned}$$

By Lemma (3),

$$\Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\} \leq \exp(n\ g(\varepsilon_a, \lambda)) \leq \exp(n\ g(\varepsilon_a, \varepsilon_a)) < \frac{\delta}{2}.$$

Hence,

$$\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r\lambda\right\} < \frac{\delta}{2} < \delta.$$

In Case (ii), we have $\Pr\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon_a\} = \Pr\{\widehat{\boldsymbol{\lambda}} = 0\}$ and

$$\begin{aligned}\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r\lambda\right\} &= \Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a\right\} \\ &= \Pr\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon_a\} + \Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\} \\ &= \Pr\{\widehat{\boldsymbol{\lambda}} = 0\} + \Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\}.\end{aligned}$$

Noting that $\ln 2 < 1$, we can show that $-\varepsilon_a < g(\varepsilon_a, \varepsilon_a)$ and hence

$$\begin{aligned}
\Pr\{\widehat{\boldsymbol{\lambda}} = 0\} &= \Pr\{X_i = 0,\ i = 1, \cdots, n\} \\
&= [\Pr\{X = 0\}]^n \\
&= e^{-n\lambda} \\
&= e^{-n\,\varepsilon_a} \\
&< \exp(n\, g(\varepsilon_a, \varepsilon_a)) \\
&< \exp\left(n\, g\left(\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) < \frac{\delta}{2}
\end{aligned}$$

where the second inequality follows from Lemma (3). Hence,

$$\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r \lambda\right\} < \frac{\delta}{2} < \delta.$$

In Case (iii), by Lemma (2), Lemma (3) and Lemma (4), we have

$$\begin{aligned}
\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r \lambda\right\} &= \Pr\{\widehat{\boldsymbol{\lambda}} \leq \lambda - \varepsilon_a\} + \Pr\{\widehat{\boldsymbol{\lambda}} \geq \lambda + \varepsilon_a\} \\
&\leq \exp(n\, g(-\varepsilon_a, \lambda)) + \exp(n\, g(\varepsilon_a, \lambda)) \\
&< \exp\left(n\, g\left(-\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) + \exp\left(n\, g\left(\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) \\
&< 2\exp\left(n\, g\left(\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) < \delta.
\end{aligned}$$

In Case (iv), by Lemma (5), Lemma (6) and Lemma (4), we have

$$\begin{aligned}
\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r \lambda\right\} &= \Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r \lambda\right\} \\
&= \Pr\{\widehat{\boldsymbol{\lambda}} \leq (1 - \varepsilon_r)\lambda\} + \Pr\{\widehat{\boldsymbol{\lambda}} \geq (1 + \varepsilon_r)\lambda\} \\
&\leq \exp(n\, g(-\varepsilon_r \lambda, \lambda)) + \exp(n\, g(\varepsilon_r \lambda, \lambda)) \\
&< \exp\left(n\, g\left(-\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) + \exp\left(n\, g\left(\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) \\
&< 2\exp\left(n\, g\left(\varepsilon_a, \frac{\varepsilon_a}{\varepsilon_r}\right)\right) < \delta.
\end{aligned}$$

Therefore, we have shown $\Pr\left\{\left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_a \ \&\ \left|\widehat{\boldsymbol{\lambda}} - \lambda\right| \geq \varepsilon_r \lambda\right\} < \delta$ for all cases. This completes the proof of Theorem 1.

References

[1] X. Chen, "Exact computation of minimum sample size for estimation of Poisson parameters," arXiv:0707.2116v1 [math.ST], July 2007.

[2] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] M. M. Desu and D. Raghavarao, *Sample Size Methodology*, Academic Press, 1990.